THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Research Data MANTRA: A Labour of Love

Citation for published version:
Rice, R 2014, 'Research Data MANTRA: A Labour of Love', *Journal of eScience Librarianship*, vol. 3, no. 1, 4. https://doi.org/10.7191/jeslib.2014.1056

Digital Object Identifier (DOI):
10.7191/jeslib.2014.1056

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of eScience Librarianship

Publisher Rights Statement:
This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

OPEN ACCESS

Download date: 07. May. 2023

Journal of eScience Librarianship

Volume 3 | Issue 1 Article 4

2014

Research Data MANTRA: A Labour of Love

Robin Rice
University of Edinburgh, r.rice@ed.ac.uk

Follow this and additional works at: http://escholarship.umassmed.edu/jeslib

 Part of the Scholarly Communication Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.

Recommended Citation
Rice, Robin. (2014). "Research Data MANTRA: A Labour of Love." *Journal of eScience Librarianship* 3(1): Article 4. http://dx.doi.org/10.7191/jeslib.2014.1056

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in Journal of eScience Librarianship by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

JESLIB 2014; 3(1): 44-51
doi:10.7191/jeslib.2014.1056

Journal of eScience Librarianship
putting the pieces together: theory and practice

Research Data MANTRA: A Labour of Love

Robin Rice

University of Edinburgh, Edinburgh, Scotland

Abstract

Research Data MANTRA (or Management Training) is a labour of love. It has been an integral part of the University of Edinburgh's Research Data Management (RDM) programme since 2012. The staff at EDINA and Data Library at the University of Edinburgh has been curating this resource, based on internal and external feedback and has just published its fourth release since 2011. MANTRA is an open, web-based training course intended for self-paced learning by PhD students and early career researchers or others who manage digital data as part of a research project. It informs about good practice in research data management with real life stories. MANTRA's approach is to be fun, relevant, useful, interactive and timely (FRUIT!). Librarians' training needs are catered for through a companion resource, the DIY RDM Training Kit for Librarians.

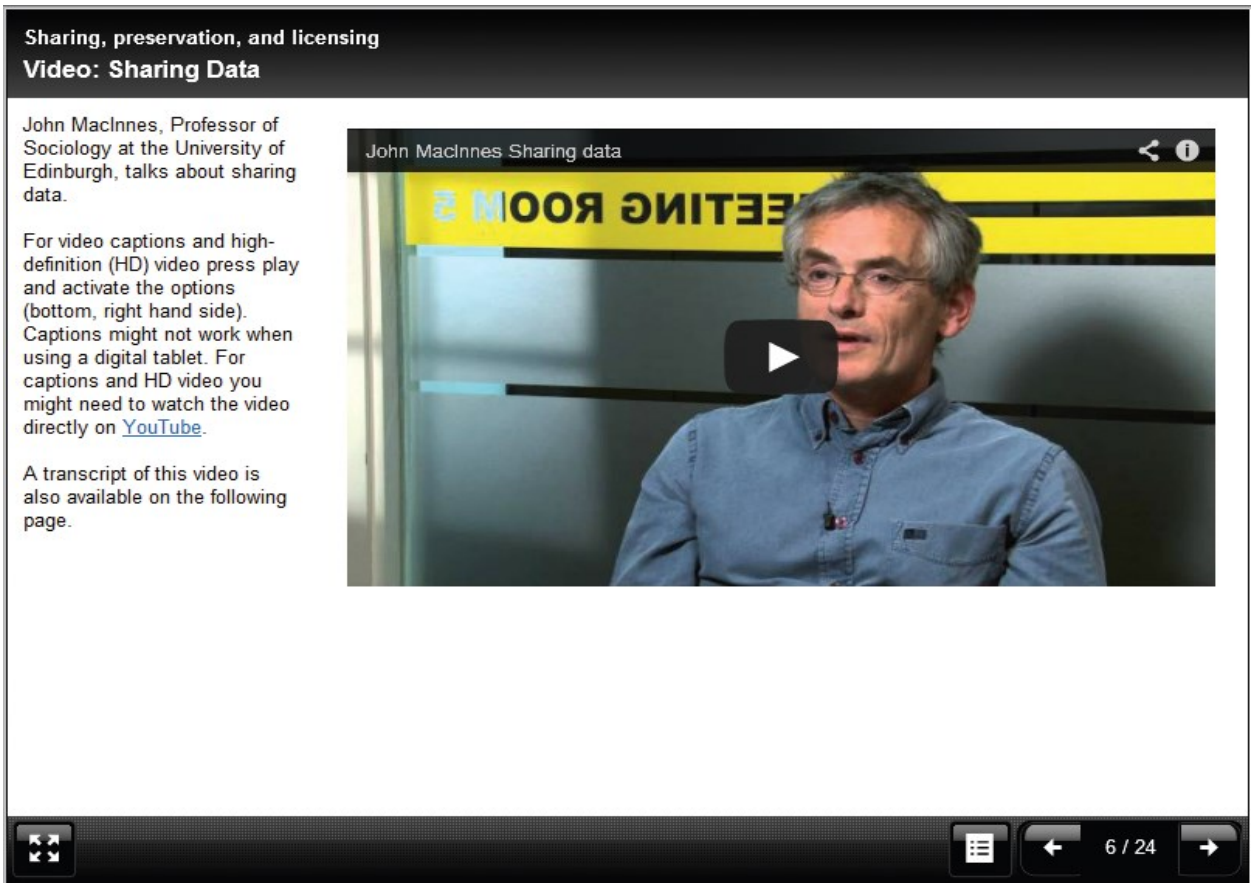
Introduction

The Research Data MANTRA (Management Training) online course is an integral part of the University of Edinburgh's Research Data Management (RDM) programme delivered by Information Services. Since it was created in 2011, the University of Edinburgh's Institute for Academic Development (IAD) has helped us introduce it to PhD students and early career researchers who have valued the course. The website is managed by EDINA and Data Library, a division of Information Services, with both national and local remit, extending into UK higher and further education and beyond. Thanks to a continuous positive feedback loop from users and web logs, this resource has avoided the fate of some completed project outputs—slow decay and eventual abandonment. To the contrary, the Data Library team continues to nurture and develop content and sustain maintenance.

Now that Edinburgh University is pursuing a data-driven research agenda through the advent of Edinburgh Data Science (Bayne, et al. 2014), it is more important than ever that we support our students and researchers to attain good practice in their data analysis and data management. Because MANTRA is openly licensed and freely available, there is no limit to who may benefit. Our usage statistics are limited due to the open nature of the course (no registration, no assessment), but Google Analytics has recorded nearly 10,000 new sessions from people in 144 countries over the last academic year, and 16,000 total sessions (Rice, 2014).

Correspondence to Robin Rice: r.rice@ed.ac.uk
Keywords: research data management, training, professional development, open educational resource, data skills

Figure 1



What is MANTRA?

The Research Data MANTRA course is a free online training course intended for self-paced learning by PhD students and early career researchers or others planning to manage digital data as part of the research process. It is available online at http://datalib.edina.ac.uk/mantra or for download as an open educational resource (OER) from the Jorum repository, http://find.jorum.ac.uk/resources/18281. It was developed at EDINA and Data Library, University of Edinburgh, by two data librarians and an e-learning specialist, with funding from Jisc in 2010-11, under the Managing Research Data Programme's training strand (Jisc, 2010).

The course aims to increase researchers' awareness of data management issues. It informs about good practice in research with real life stories and scenarios. The eight online learning modules take an hour or less to complete, and are designed to be fun, relevant, useful, interactive, and timely (FRUIT). For example, images are used to break up text, short video clips of researchers talking about their own experience appear in each unit (Figure 1), factual scenarios encountered by researchers in different disciplines are conveyed, and interactive quizzes about the content is presented at regular intervals.

Accompanying PDF tutorials give a deeper grounding in data handling skills, such as importing and exporting data, recoding, and keeping track of transformations within four common software packages using open, empirical datasets, available for download from the website. The analysis environments covered are NVivo, SPSS, ArcGIS, and R

(packages identified during the original needs assessments with project stakeholders).

What Are the Topics Covered?

The online module contains eight learning units:

- Research data explained: the concept of research data, what constitutes research data, and how it differs from other types of information; the data life-cycle, big data, and data-driven research.

- Data management plans: good data management, research funder requirements, how to create a data management plan; an editable, downloadable template for a postgraduate research project.

- Organising data: organisation or 'housekeeping,' why it is important from the start of a research project; what constitutes good data file management, such as file naming and versioning.

- File formats and transformation: open vs proprietary data file formats, compression, normalisation, and other kinds of data transformations; why such skills are useful in a research context.

- Documentation, metadata, citation: the importance of documenting data during a research project, re-using other people's data, purposes of metadata; the importance of proper data citation and reproducible research.

- Storage and security: issues involved in storing, securing and backing up research data. Storage media, checking your back-ups, cloud storage, password safety, encryption, destroying sensitive data.

- Data protection, rights and access: ethical obligations towards human subjects, confidentiality, privacy, confidentiality breaches in the news, legal requirements for personal and sensitive data; intellectual property rights; anonymization techniques; the public's rights to access publicly funded research under freedom of information law.

- Preservation, sharing and licensing: benefits and challenges of sharing data, drivers for sharing, scientific fraud, risks to the longevity of digital data, the importance of digital preservation and trusted repositories, open data and linked open data, open data licenses, Creative Commons 4.0 and CC-Zero.

Each unit ends with an interactive summary to reinforce learning and suggests 'Next steps' with 'Further reading' on each topic. A video playlist for each unit is available separately on the Data Library Youtube channel: www.youtube.com/user/maltsvid.

The online module was created using the Xerte Online Toolkits (XOT) open-source authoring tool developed by the University of Nottingham. XOT allows for the creation of content using page templates that present material in a variety of ways (such as text plus image or embedded video). Standard-compliant (SCORM) learning packages have been exported for deposit in Jorum, which others can access and import to any virtual learning environment. The Creative Commons Open Attribution licence ensures maximal flexibility for those wishing to repurpose and rebrand the materials for use in their own institutions.

Why 'MANTRA'?

A new colleague recently asked why we call this training course MANTRA. When we were first funded by Jisc to produce the training materials in 2010-11, it seemed a useful acronym to avoid the term 'RDM.' The more we thought about it, the more apt it seemed to be. Focusing on students and early career researchers, MANTRA is about

trying to create culture change—with or without the big sticks of the funders' mandates. If we talk about good practice in data management often and loud enough, it may happen (like a mantra).

An academic at our institution once challenged our use of the colourful image on our flyer (http://edin.ac/1pgpTop), seen partially on the home page banner—of a Tibetan monk's hands carefully adding the last grains of sand to a beautiful sand mandala. She said the purpose of creating a sand mandala is so it can be destroyed, symbolising lack of attachment—not exactly a symbol of sustainability. Her point was well-taken, but we felt that our use of this image can symbolise how easily one's data can be lost if precautions are not taken. It has happened to everyone on a small or grand scale. Moreover, the hands of the monk can symbolise the actions of the data curator to care for the dataset.

The Origins of MANTRA

Originally, we were funded to produce training for PhD students: "JISC Call for Projects 04/10 sought to fund projects to encourage research data management training to be embedded in post-graduate academic curricula" (Jisc, 2013). We had already done a pilot RDM training course with some GeoSciences doctoral students in December, 2009 and felt we could apply the experience (both positive and negative outcomes) to a larger project.

Our project partner, the Institute for Academic Development, convinced us to a) run the course as a stand-alone online course using the model of *PG Essentials*, purchased from Melbourne University containing optional, self-paced online modules such as 'writing a literature review' that appeal to new PhD students for their pragmatism; and b) to include early career researchers (aka post-docs), as they too are working on their research skills and are within the target audience of IAD. This seemed entirely reasonable to us, since it is staff that faces the greatest culture change with regard to RDM requirements and expectations.

Although we were funded to create discipline-specific training materials, we really wanted to try to create something compelling across a broad range of disciplines (or 'schools' at our University—the equivalent of academic department). So we targeted three schools with post-graduate programmes, and created one resource for all three, in the hope that they would be representative enough to work for others as well. The programmes were Social and Political Studies (SPS)—which had both quantitative and qualitative data analysis requirements: the Doctorate in Clinical Psychology—which was unusual in that much of the programme was offered as distance education, and students spent much time in the field with National Health Service staff supervisors; and GeoSciences—with whom, as with SPS, we had much prior contact. The three schools happily spanned the overarching colleges of the University. All shared a concern with managing confidential personal data. We felt—and still do, confirmed by anecdotal feedback—that we had at least covered the ground of disciplines where there is a 'lone researcher,' if not research groups using sophisticated computing systems.

Before writing the learning materials, we conducted a needs assessment with PhD training managers in each school based on a single recorded face-to-face interview with each. The IAD director had advised us not to expect too much time from these project stakeholders, so we were careful to try to get what we needed from one interview (Rice, 2010). Over the one-year project, schools experienced a high amount of academic staff turnover—maternity leave, sabbaticals, promotions and exits—making it difficult for us to meet their successors and encourage follow-through on their end. Each school managed to 'embed' MANTRA into its training curriculum either through induction and training activities or in the required coursework,

although it has remained an optional activity, even in required courses. In the GeoSciences, we have been invited by one of the successors into the classroom to introduce MANTRA and help the students draft a Data Management Plan (DMP) for their research proposals. However, we have not been involved in evaluating the DMPs.

The school-based stakeholders also helped us identify the most useful analysis environments on which to focus the data handling tutorials for them and agree to the layout and learning objectives with each of the commissioned authors (experts in each software package). The idea of the data handling tutorials came from the Data Library and was included in the original funding bid as a 'unique selling point.' Our experience with the pilot workshop indicated students wanted hands-on exercises as well as being told general principles. As eager new researchers, they tend to be more interested in data analysis than data management, and our experience as data librarians told us that data handling could be a bridging skill facilitating both. For example, to conduct statistical analysis on a large dataset, certain preparation is needed, especially on other people's data. Moreover, the Data Library led an earlier project aimed at bringing 'real data' into learning environments to raise students' quantitative skills, particularly in the social sciences, so this was a way to further that aim as well (Rice and Fairgrieve, 2003).

One finding that emerged from the needs assessment was that unlike ourselves, the doctoral trainers (at that time) were not concerned with teaching their students to share or preserve their data at all, never mind how to do this well. This may have been partly due to the timeframe (the funders' requirements that have driven awareness were still emerging) and partly due to the bias that a novice's data is not worth sharing or preserving. For this reason, the final two units of MANTRA were 'shelved' for a second release and were not finished during the course of the funded project. However, they were both released in a second 2012 version as we viewed the topics as crucial to the spirit of RDM. As with the first six units, we conducted user testing with doctoral students from the three schools to ensure the usability and usefulness of the new material.

An external evaluation of the MANTRA project was completed in 2012 by a staff member of the Digital Curation Centre based at the University of Glasgow, after the course had been available for one academic year (Molloy, 2012). By this time MANTRA had developed a good reputation in the emerging 'RDM community.' However, evidence was limited by our basic web statistics, in that we did not have the full tracking capability of user behaviour that a Virtual Learning Environment could have captured. (Google 'universal' analytics with event tracking may provide new insights over the coming year.)

Lessons Learned for the Third (2013) Release

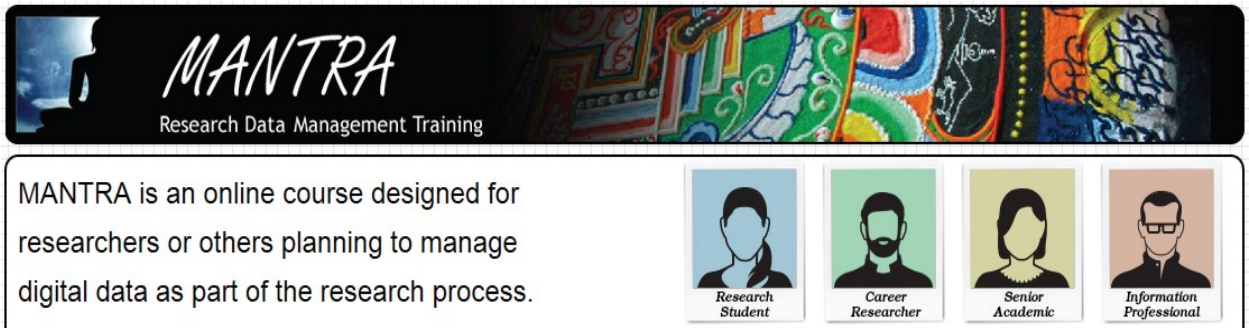
The third, 2013 release combined minor cumulative edits to the online module, a major change in the presentation of the course via the home page - done to our specification by a web and graphic designer, and a back-end upgrade to the XOT software which resulted in html5-compliant web pages. The significance of moving to this standard and away from Flash meant that the units could be viewed on iPads for the first time: a significant user requirement for us.

Since our University's RDM Policy was passed (2011) and an RDM Roadmap published (2012), the RDM Steering Group that governed the programme saw fit to test MANTRA across the University. However, they and their colleagues were mainly in the College of Sciences and Engineering, particularly Physics and Informatics, so there was some concern that the first unit, *Research data explained*, was too basic (the Big Data section had not yet been written).

We also had feedback from the Institute for

JESLIB 2014; 3(1): 44-51
doi:10.7191/jeslib.2014.1056

Figure 2



Academic Development (IAD) staff that learners need to know what they're going to get out of a resource before they invest their time in it. So the third, 2013 release of MANTRA included four personas in the top right corner (Figure 2). If a 'career researcher' clicks on their persona, the page suggests that they skip over the first unit and begin with *Data Management Plans*. 'Research students' can start with the first unit, whereas 'senior academics' are shown the *Data handling tutorials* for their students, and to encourage culture change, the *Sharing, preservation and licensing* unit. We also added some JavaScript that gives a peek at each unit's content without leaving the home page and tried to make the data handling tutorials stand out by using a different background colour on the button.

We suggest that information professionals start with *Data protection, rights and access*, moving onto *Data sharing and preservation*, since those are likely to be the topics they need to know best to give staff and students RDM support and advice. But we also have a special offering for librarians, described below.

The DIY RDM Training Kit for Librarians

To help professional librarians grasp supporting RDM, the DIY Research Data Management Training Kit for Librarians is a companion to MANTRA that will give small groups everything they need to conduct a RDM training. We have used the materials in this kit to train all of our academic service librarians, co-facilitated by a data librarian and an academic service librarian. The idea is that with the help of a good facilitator (i.e., someone who is able to stimulate discussion and involve the entire group), the training can be done without an 'expert' instructor. In fact, the facilitator can be one of the learners, or the role could even be shared amongst the participants.

The training comprises five two-hour, face-to-face sessions, with learning and preparation time between each session of two weeks or more. The sessions open with short talks (from local speakers or alternatively, podcasts from the kit) followed by group exercises created by the UK Data Archive, with lots of supportive, small group discussion. MANTRA modules are used as reading assignments and accompanying 'reflective questions' are used to help librarians 'put themselves in the shoes of the researcher.' Learning is reinforced and put into practice through an independent study assignment of completing and publishing an interview with a researcher, resulting in a simple data curation profile—ideally, to be publicly shared. The independent study is optional, since the target group is busy professionals, though we feel it is crucial for building confidence needed to take the training into practice, and also to appreciate the anxiety many researchers feel about being 'forced' to share their own data online.

There are five topics, loosely coupled with MANTRA's eight units:

- Data management planning
- Organising & documenting data
- Data storage & security
- Ethics & copyright
- Data sharing

The DIY Training Kit complements other librarian training resources that have emerged in the last couple of years. Some of these were compared in context in a paper written for a DigCurV (Digital Curator Vocational Education Europe) conference (Macdonald & Rice, 2013, 3). Anecdotal evidence suggests the kit has succeeded outside of the University of Edinburgh librarian training, both in the UK and the US. Lisa Haddow has posted on the Edinburgh Data Blog describing Stirling University's use of the course for training its subject librarians (Haddow, 2014).

The Fourth Release: 2014

September 5, 2014 marked the fourth release of MANTRA. The data handling tutorials required an update for newer software versions and better usability, and a number of newer concepts had become prevalent in the field of RDM, such as data citation; Creative Commons 4.0 was now compatible for research data, the University's Records Manager had volunteered some suggestions about the Data Protection Act. It was time to do a semi-major overhaul of the written content of the online modules, in order to keep MANTRA from fading into obsolescence.

Data Library team members have edited additional video material, improved the language and flow of the units, suggested new content and references, and helped test, revamp, and update the data handling practical exercises (our small team has experienced some flux as well as growth thanks to the University's RDM programme). A change control document was used as we updated the online modules, partly to give each unit's author control over suggested edits and partly for those who have already adapted a previous version in a local VLE to be able to easily view the new and edited content. As the structure has not been altered, it should be a smooth experience for returning users and those who have recommended the course to students based on earlier versions. Three remote usability tests gave us ideas for minor enhancements to the wording and navigation of the home page. We also upgraded our Creative Commons license to 4.0 and added standard cookie and website accessibility statements in the footers.

Another change at Edinburgh this year is that together with colleagues in Information Services we are offering a range of face to face RDM training courses, some for postgraduates and some for staff (http://edin.ac/1usi2JP). Some of these use MANTRA as a basis, including the data handling practicals.

Conclusion

We were fortunate to have been early on the scene with MANTRA. Although we had committed ourselves to open access in our funding bid, it became a requirement for all of the projects in the Jisc Managing Research Data (MRD) training strand once funded. All of the other projects resulted in discipline-specific materials such as PowerPoint files used in traditionally taught courses, so MANTRA was the most broadly accessible outside of our own institution, and was later taken up by several of the institutions funded by successive Jisc MRD programmes.

Research Data MANTRA has seen four releases over as many years, and has attracted a number of fans in the UK and beyond. We know that it is used in regular postgraduate training at Oxford University, that Melbourne University is adapting it for its own postgraduate training, that the University of Montreal plans to translate it into French,

and Minho University has expressed interest in translating it into Portuguese. Anecdotally, librarians have expressed gratitude that it has saved them from developing their own researcher or librarian training from scratch. It has been a labour of love because the team has remained dedicated to its quality assurance, and to keep the course openly licensed and free to all comers.

References

Bayne, Sian, Kevin Collins, Elizabeth Elliot, Richard Kenway, Andrew Morris, and Dave Roberts. 2014. "*Edinburgh Data Science: Working with Data to Support Research, Innovation, and Education.*" [discussion paper]. Edinburgh: University of Edinburgh. http://www.datascience.ed.ac.uk/sites/default/files/attachments/EDSfinal.pdf

Haddow, Lisa, 2014. "Training subject librarians in RDM." *Edinburgh Data* (blog), January 9. http://datablog.is.ed.ac.uk/2014/01/09/training-subject-librarians-in-rdm/

Jisc. 2010. "Research Data Management Training Materials (RDMTrain)." Last modified 20 November, 2013. http://www.jisc.ac.uk/whatwedo/programmes/mrd/rdmtrain.aspx

Macdonald, Stuart, and Robin Rice. 2013. "'DIY' Research Data Management Training Kit for Librarians." In *Proceedings of the Framing the Digital Curation Curriculum conference*. Florence, Italy, 6-7 May. http://digcur-education.org/eng/Resources/DigCurV-2013-proceedings/MacDonald-paper27

Molloy, Laura. 2012. "JISC Research Data MANTRA project at EDINA, Information Services, University of Edinburgh: Evaluation." Glasgow: University of Glasgow. http://eprints.gla.ac.uk/71435/1/71435.pdf

Rice, Robin. 2010. "Notes from Needs Assessment with Doctoral Programmes." Available from: Research Data MANTRA Project: Deliverables. Last updated July 31, 2013. http://www.ed.ac.uk/schools-departments/information-services/about/organisation/edl/data-library-projects/mantra/deliverables

Rice, Robin. 2014. "New Release of Research Data MANTRA (Management Training) Online Course." *Edinburgh Research Data Blog,* September 5. http://datablog.is.ed.ac.uk/2014/09/05/new-release-of-research-data-mantra-management-training-online-course/

Rice, Robin, and Joan Fairgrieve. 2003. "An Enquiry into the Use of Numeric Data in Learning and Teaching in UKHE." *Radical Statistics.* 81:18-33. London: Radical Statistics. http://www.radstats.org.uk/ no081/rice.pdf

Disclosure: The author reports no conflicts of interest.

All content in Journal of eScience Librarianship, unless otherwise noted,
is licensed under a Creative Commons Attribution-Noncommercial-Share Alike License
http://creativecommons.org/licenses/by-nc-sa/3.0/

ISSN 2161-3974